Supporting Information (SI): Transient Pupil Constriction Reflects and Affects Facial Attractiveness

Three experiments (SE1-3) were conducted to examine (1) the potential factors accounting for the inconsistency between our finding (pupil constricts to attractiveness) and the literature which demonstrated pupil dilation to attractiveness, and (2) whether the attractiveness judgment for geometric figures could indeed not induce corresponding pupil constriction. The potential factors included stimulus presentation time (until a response, usually about 2 s, or a response postponed until 5 s), task demand (attractiveness judgment or passive viewing), stimulus category (faces or geometric figures), and sequential contrast induced by luminance change (pupil constriction or dilation, in general, in response to sequential luminance change). The overall results indicate that none of the above factors alone can explain the discrepancy between our finding and the literature. However, in general, the effect of pupil constriction to attractive faces (still not to geometric figures) was more effectively observed during the early time course (about 2 s after the stimulus presentation, SE1 but not SE2) when the pupil generally constricted in response to the relative sequential luminance increase (SE3) and when a task demand was required (SE1 but not SE2). Pupil dilation to attractiveness was occasionally observed only for the face rated most attractive (i.e., 8 or 9 rating score), either during a later time period (3 to 5 s after stimulus onset) or passive viewing (SE2).

Supplementary Experiment 1

Experiments 1 and 2 demonstrated that the pupil constricted to attractive faces within 2 s after the stimulus onset, whereas it is unclear how the effect went after 2 s.

To address the issue, we presented the face for 5 s and asked participants to hold their judgment response until the face disappeared. Their pupillary responses were recorded during the whole 5-s inspection period.

Materials and Methods. The same group of participants in Experiment 4b (11 females, age range of 20-48, median age = 40 years) participated in this experiment. The stimuli and experimental procedure were the same as in Experiment 2, except for the following. First, only one condition, attractiveness judgment for faces, was conducted. Second, the face was presented on the screen for 5 s and then replaced by the fixation display. Participants were asked to give the rating score after the face disappeared.

Results. The blink rate during the -1–to 5-s time window reference to the target onset was 11.7%. Pupil response results are shown in Fig. SE1. Mean pupil diameter 0.5–1.5 and 3–5 s after the target onset was subjected to a repeated-measure ANOVA, as was done in Experiments 1 and 2. Results of the ANOVA and linear trend analysis are shown in Supplementary Table 2. Pupil constriction to attractive faces was found not only 0.5–1.5 s ($p < .001$) but also 3 to 5 s ($p < .01$) after the stimulus onset.

Supplementary Experiment 2

This experiment aimed to re-examine the effect of pupil constriction to attractive faces over time by presenting the stimuli with less luminance/contrast change. We also added a passive viewing condition to examine the effect of task demand. We used line-drawing faces which were presented with their luminance contrast enhanced gradually for 1 s, stayed on the screen for 3 s, and then disappeared

gradually for 1 s. In the attractiveness judgment condition, participants evaluated the attractiveness of the face and gave their answer after the 5-s stimulus presentation, as was done in Supplementary Experiment 1. In the passive viewing condition, participants viewed the same set of images without any task demand. The two conditions were conducted in counterbalanced order across participants. Pupil data in the passive viewing condition were sorted by attractiveness ratings obtained in the attractiveness judgment condition for each participant individually.

Materials and Methods. A different group of twelve adults (9 females, age range of 21-40, median age = 35 years) from the main experiments participated in this experiment. Forty face images (five in each subcategory as used in Experiment 1) were selected for further processing. Each face was manually line-drawn based on the original image with the pupil adjusted in five different sizes (see example in Fig. SE2-1). The face was presented at the center of the screen against a gray background (21.04 cd/m^2). The interlaid fixation display was the same as in Experiment 1. We applied the nested design in which each participant viewed 40 face images only, all with different identities. Each pupil size level was presented eight times with different face identities. There were two conditions: attractiveness rating and passive viewing. The two types of task demand were conducted as within-subject factors in separate blocks with the order counterbalanced across the participant. The same set of stimuli was used in the two conditions, with the trial order randomly assigned. The face image was presented with its luminance contrast gradually increased for 1 s, stayed at the center of the screen for 3 s, and then gradually disappeared for 1 s. In the attractiveness rating condition, participants were asked to hold their response during the 5-s stimulus presentation and respond after the target image disappeared. In the passive viewing condition, they were asked to look at the faces without any task

involved. After the participants made the response or 0.5 s after the stimulus presentation in the passive condition, the next trial started with the 3-s fixation display. Pupillary responses were recorded throughout the experiment, as was done in the other experiments.

Results. The blink rate during the -1–5-s time window reference to the target onset was 21.4%. Pupil response results are shown in Fig. SE2-2. The same analysis as in Supplementary Experiment 1 was conducted. Results are shown in Fig. SE2-2 and Supplementary Table 2. In the attractiveness rating condition, there was a tendency of pupil constriction to attractive faces during the time course of 0.5–1.5 s after the stimulus onset, though it was not statistically significant ($p = .2$). In contrast with the result in Supplementary Experiment 1, the most attractive faces (rated as 1 or 2) induced the strongest pupil dilation response ($p < .01$) during the later time course (3–5 s). In the passive-viewing condition, although none of the effects were significant, there was a tendency of pupil dilation to the most attractive faces during the early time course ($p = .05$).

Supplementary Experiment 3

In Experiments 1 and 2, pupil constriction to attractiveness was found for faces and natural scenes but not for geometric figures. This could be due to the stimulus category or the sequential luminance contrast in which pupils in general dilated to the presentation of geometric figures (while it constricted to faces and natural scenes). To examine this issue, we manipulated the luminance of the fixation display prior to the target display so that the pupil response baseline to sequential contrast was constriction to geometric figures and dilation to faces. We examined

whether the baseline pupil response (constriction or dilation) is critical to the effect of pupil constriction to attractiveness.

Materials and Methods. Another group of ten adults (7 females, age range of 22-43, median age = 36 years) participated in this experiment. The stimuli, design, and procedure were the same as in Experiment 1, except for the following two points. First, the faces were presented at the center of the screen against a white background (94.04 cd/m^2). The interlaid fixation display for faces consisted of a light gray fixation cross (58.66 cd/m^2) against the white background. For geometric figures, the background was black (0.35 cd/m^2), and the fixation cross was dark gray (2.44 cd/m^2) in the interlaid fixation display. Second, there were only two conditions: face-attractiveness and geometric figure-attractiveness conditions (no face-roundness condition).

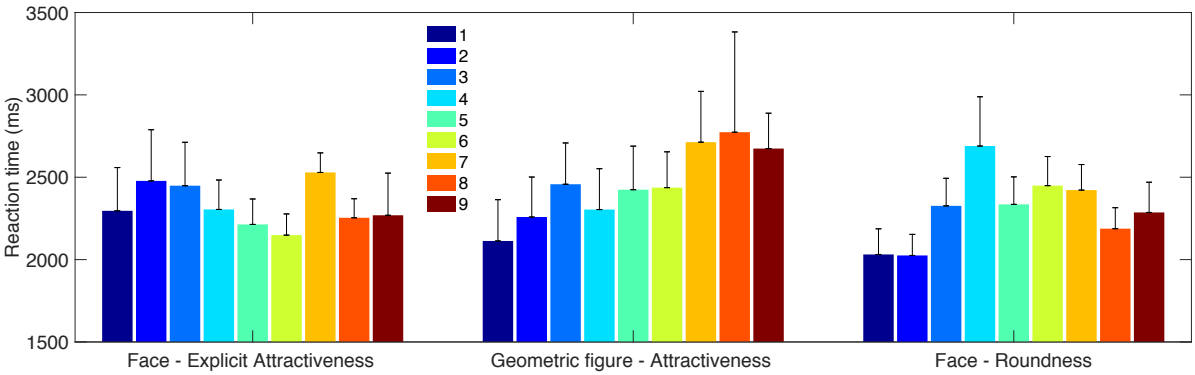
Results. The blink rate during the -1–2 s reference to the stimulus onset was 19.1%. Pupil response results are shown in Fig. SE3, and those of the statistical analysis following the same procedure as in Experiments 1 and 2 are shown in Supplementary Table 1. The amount of pupil size change did not significantly correlate with the attractiveness judgment on geometric figures nor on faces ($ps > .2$), whereas the tendency of pupil constriction to attractiveness was consistent for faces and opposite for geometric figures.

Supplementary Table 1. Statistical results for the mean pupil size during specified time window after the stimulus onset in Supplementary Experiments 1-3.

				ANOVA		Linear Trend Analysis	
	Task	Stimulus	Time (s)	*F*-statistic	*p*-value	*F*-statistic	*p*-value
E1	Att.	FC (equal lum.)	0.5~1.5	7.09	<.001***	18.98	<.001***
			3~5	4.54	<.001***	9.41	<.01**
E2	Att.	FC (line-drawing)	0.5~1.5	0.81	.57	1.69	.22
			3~5	3.10	<.01**	2.89	.12
	Pas.	FC (line-drawing)	0.5~1.5	2.21	.05	2.88	.12
			3~5	0.17	.99	0.48	.50
E3	Att.	FC	0.5~1.5	0.57	.75	1.29	.29
		GF	0.5~1.5	0.43	.86	0.82	.39

Att. = Attractiveness task, Pas. = Passive viewing, FC = faces, GF = geometric figures

(A) Mean reaction times as a function of rating scores. Error bars represent standard errors among participants.



(B) Histograms of rating scores.

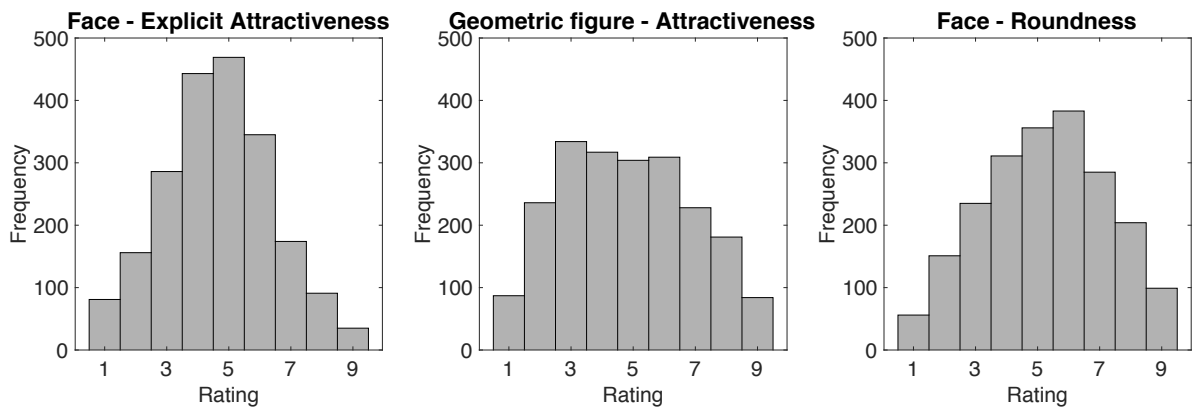
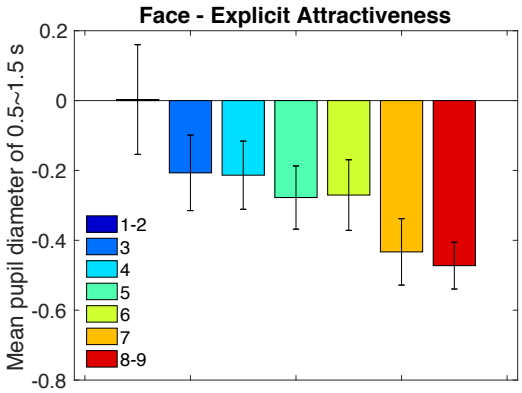


Figure S1. Supplementary information to Table 1. Behavioral results in Experiment 1.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard errors among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5-1.5 s after stimulus onset) as a function of explicit attractiveness rating for faces in Experiment 1. Each plot represents each participant. Each dot represents response to each face. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, and the darker pink (salmon-red) represents statistical significance of a positive Pearson's correlation. For this condition, ten out of 13 participants showed a tendency of negative correlation between pupil size and explicit attractiveness rating, among which six reached statistical significance individually.

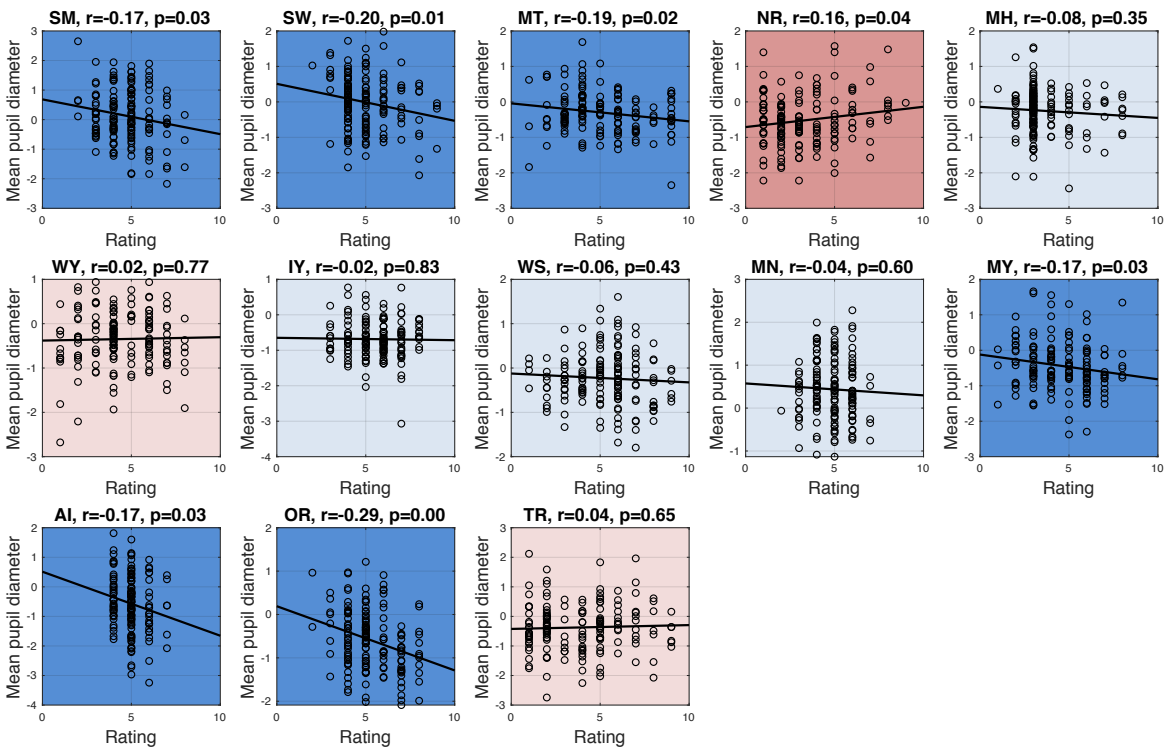
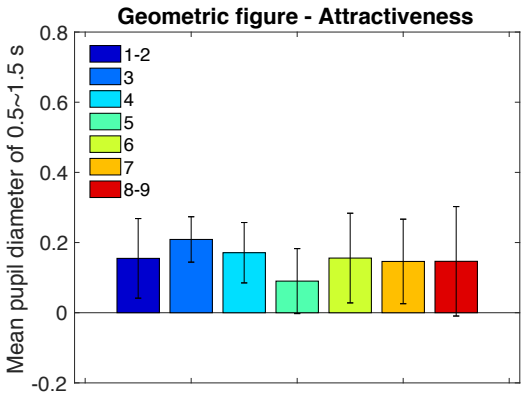


Figure S1A. Supplementary information to Figure 1(A). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the explicit facial attractiveness condition in Experiment 1.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard errors among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of attractiveness rating for geometric figures in Experiment 1. Each plot represents each participant. Each dot represents response to each geometric figure. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, and the darker pink (salmon-red) represents statistical significance of a positive Pearson's correlation. For this condition, no consistent tendency between pupil size and attractiveness rating was found among the participants (half showed positive and half showed negative).

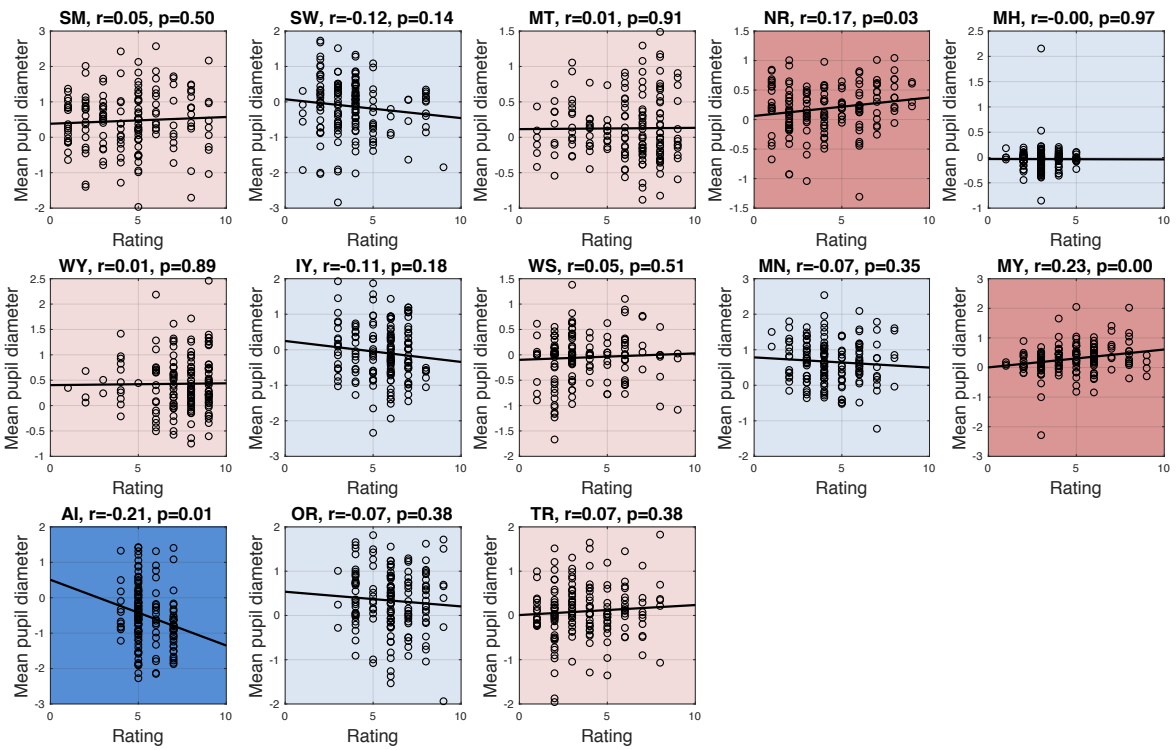
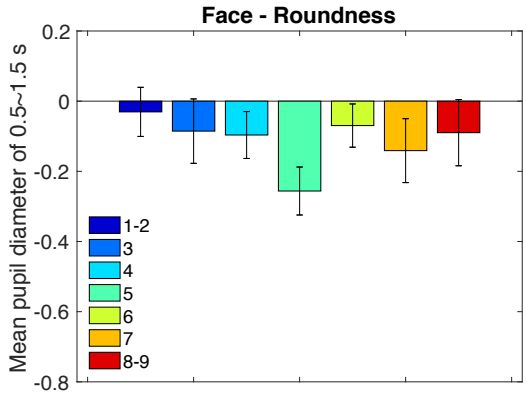


Figure S1B. Supplementary information to Figure 1(B). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the geometric figure attractiveness condition in Experiment 1.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard error among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of roundness rating for faces in Experiment 1. Each plot represents each participant. Each dot represents response to each face. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, and the darker pink (salmon-red) represents statistical significance of a positive Pearson's correlation. For this condition, no consistent tendency between pupil size and roundness rating was found among the participants.

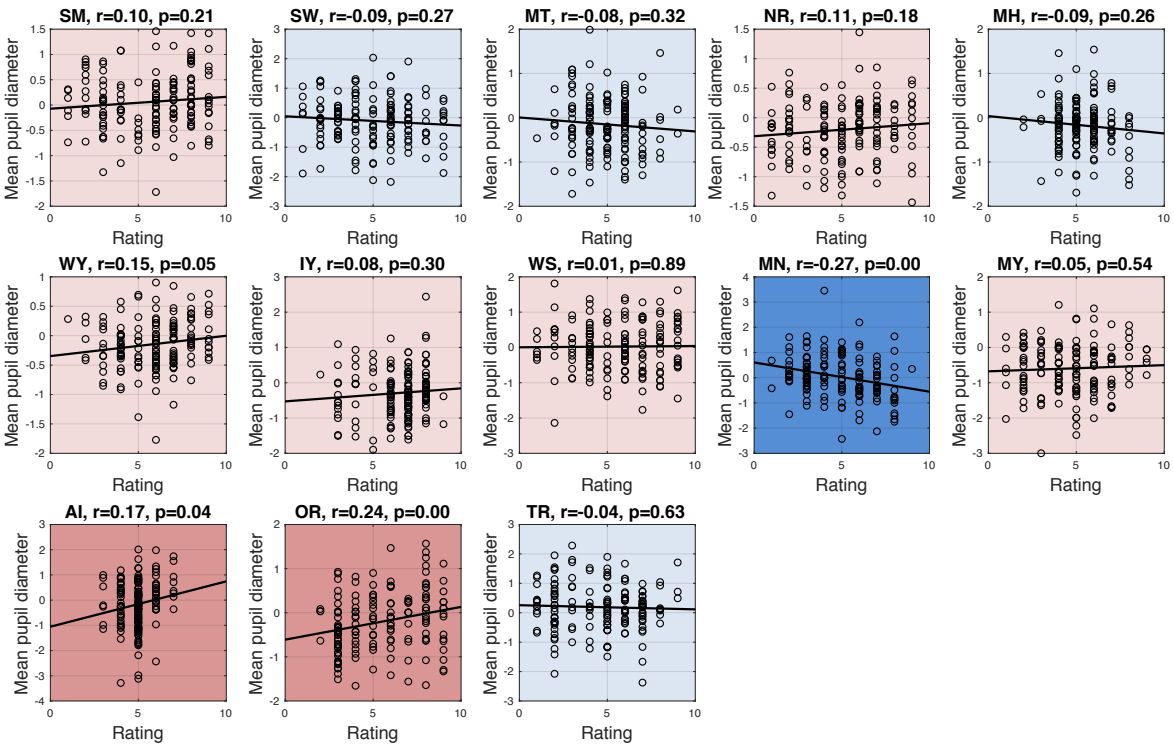
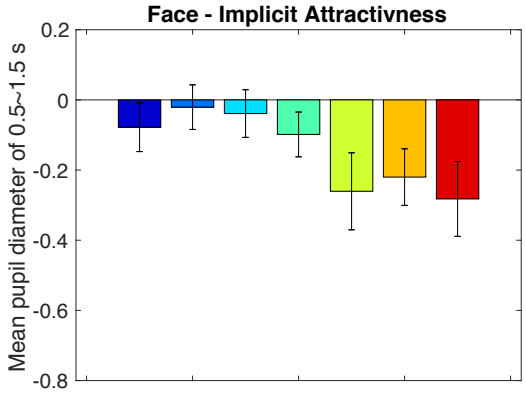


Figure S1C. Supplementary information to Figure 1(C). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the face roundness condition in Experiment 1.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard errors among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of implicit attractiveness rating for faces in Experiment 1. Each plot represents each participant. Each dot represents response to each face. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, but it did not reach statistical significance. For this condition, 12 out of 13 participants showed a tendency of negative correlation between pupil size and implicit attractiveness rating, among which four reached statistical significance individually.

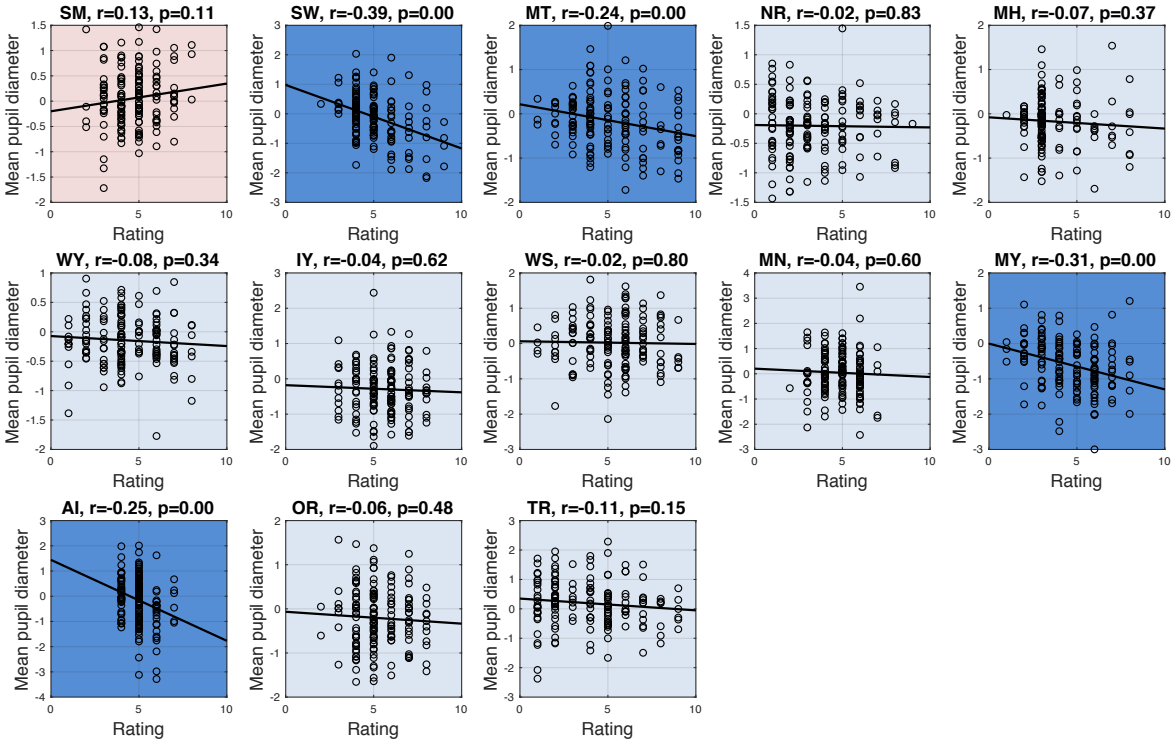
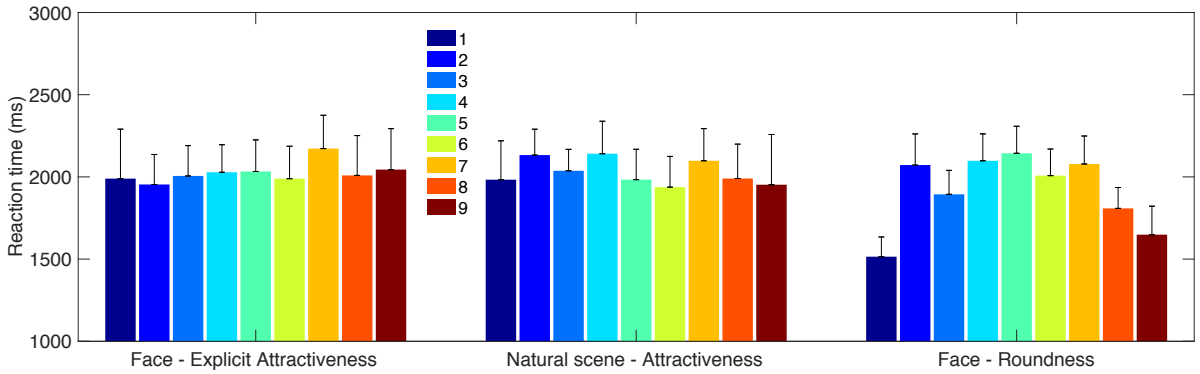


Figure S1D. Supplementary information to Figure 1(D). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the implicit facial attractiveness condition in Experiment 1.

(A) Mean reaction times as a function of rating scores. Error bars represent standard errors among participants.



(B) Histograms of rating scores.

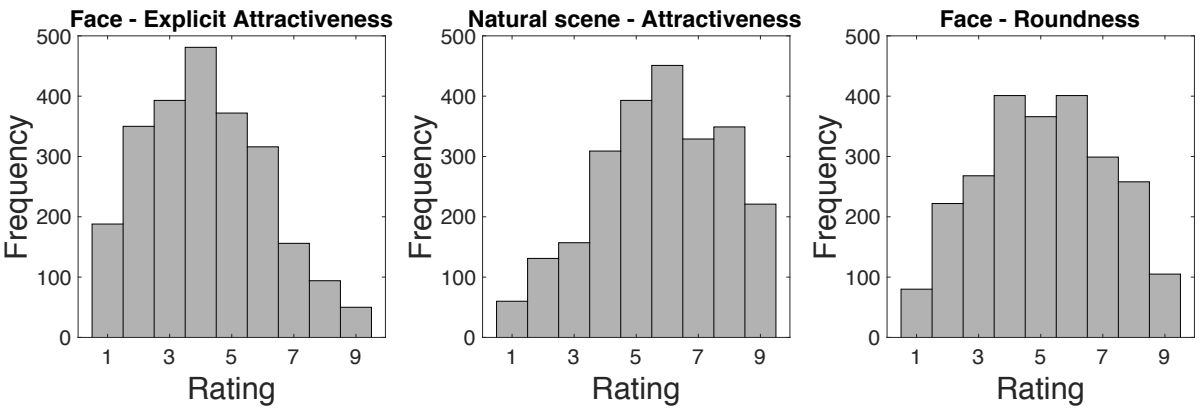
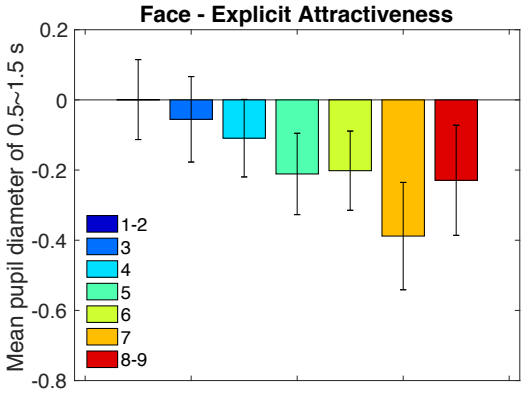


Figure S2. Supplementary information to Table 1. Behavioral results in Experiment 2.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard errors among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of explicit attractiveness rating for faces in Experiment 2. Each plot represents each participant. Each dot represents response to each face. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, and the darker pink (salmon-red) represents statistical significance of a positive Pearson's correlation. For this condition, 13 out of 15 participants showed a tendency of negative correlation between pupil size and explicit attractiveness rating, among which five reached statistical significance individually.

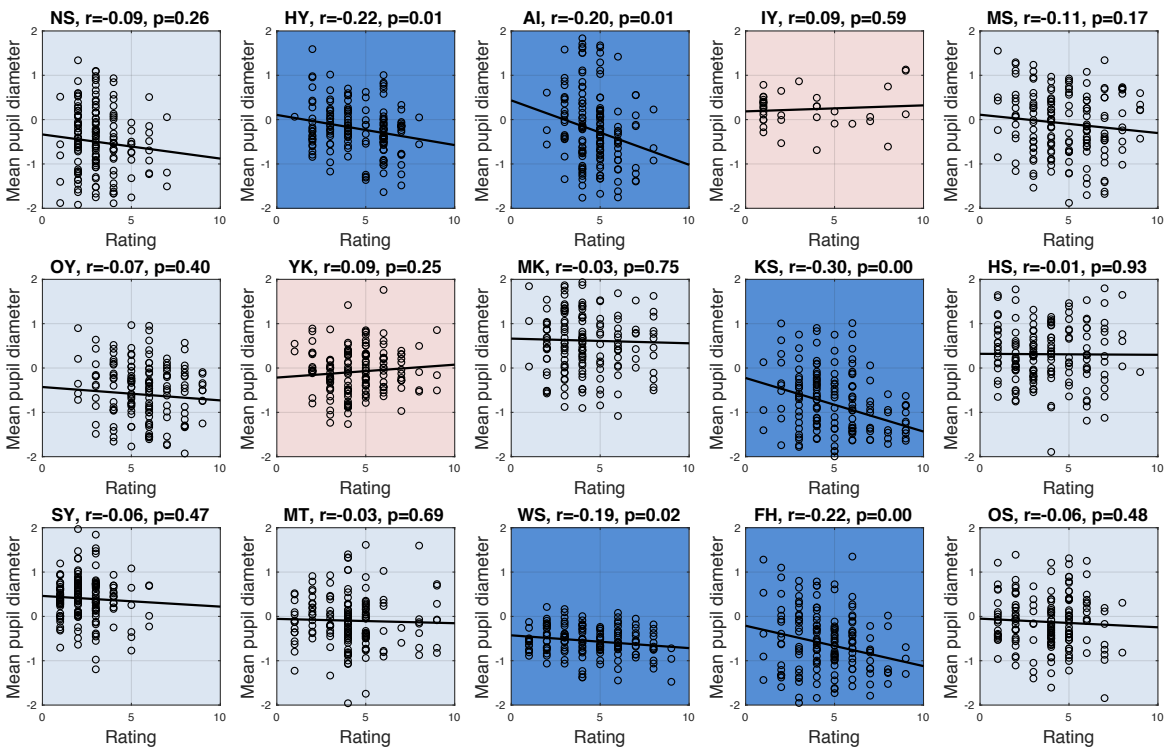
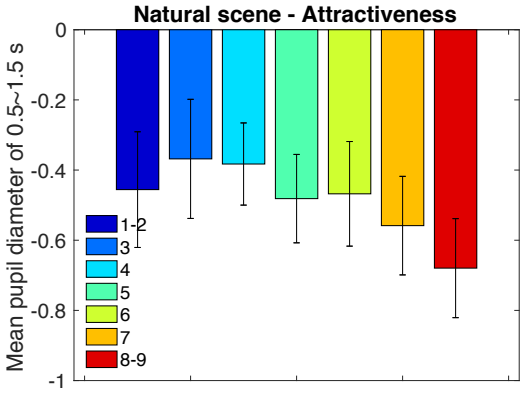


Figure S2A. Supplementary information to Figure 2(A). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participant in the explicit facial attractiveness condition in Experiment 2.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard errors among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of attractiveness rating for natural scenes in Experiment 2. Each plot represents each participant. Each dot represents response to each natural scene images. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, and the darker pink (salmon-red) represents statistical significance of a positive Pearson's correlation. For this condition, 13 out of 15 participants showed a tendency of negative correlation between pupil size and explicit attractiveness rating, among which three reached statistical significance individually.

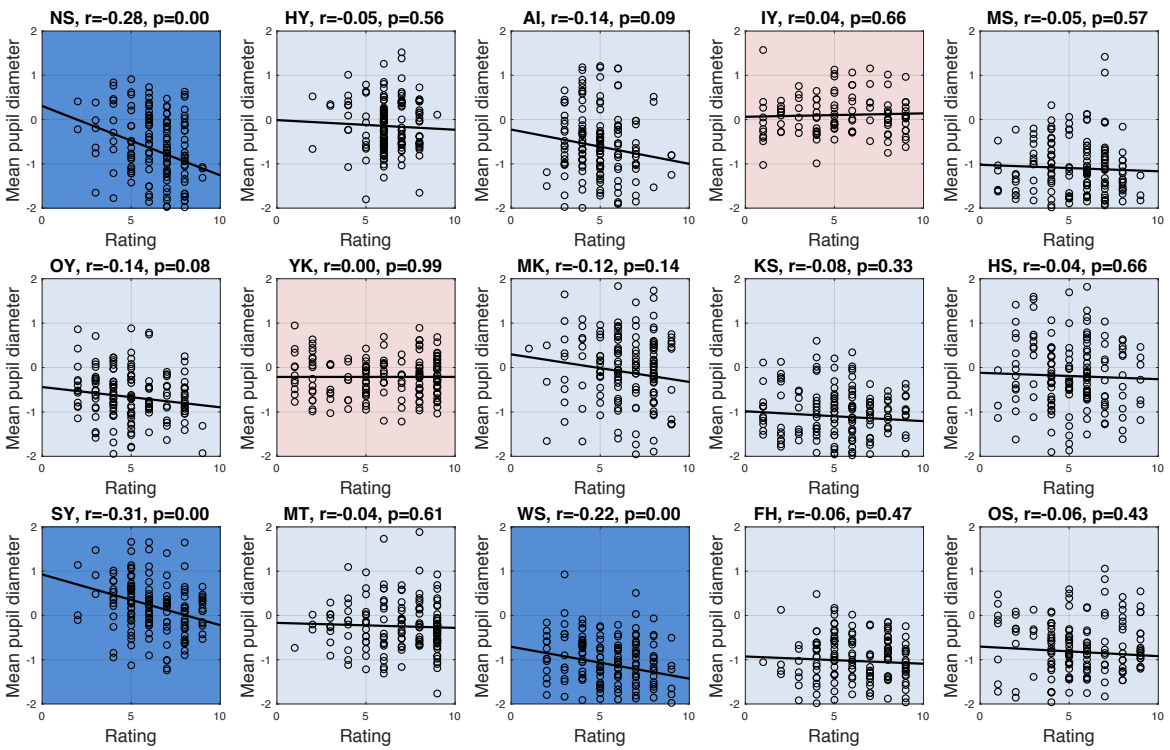
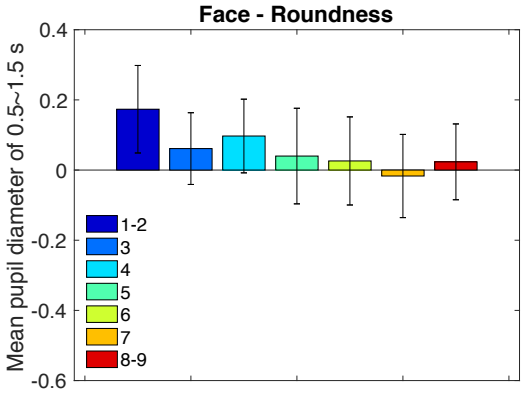


Figure S2B. Supplementary information to Figure 2(B). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the natural scene attractiveness condition in Experiment 2.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard error among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of roundness rating for faces in Experiment 2. Each plot represents each participant. Each dot represents response to each face. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, and the darker pink (salmon-red) represents statistical significance of a positive Pearson's correlation. For this condition, no consistent tendency between pupil size and roundness rating was found among the participants.

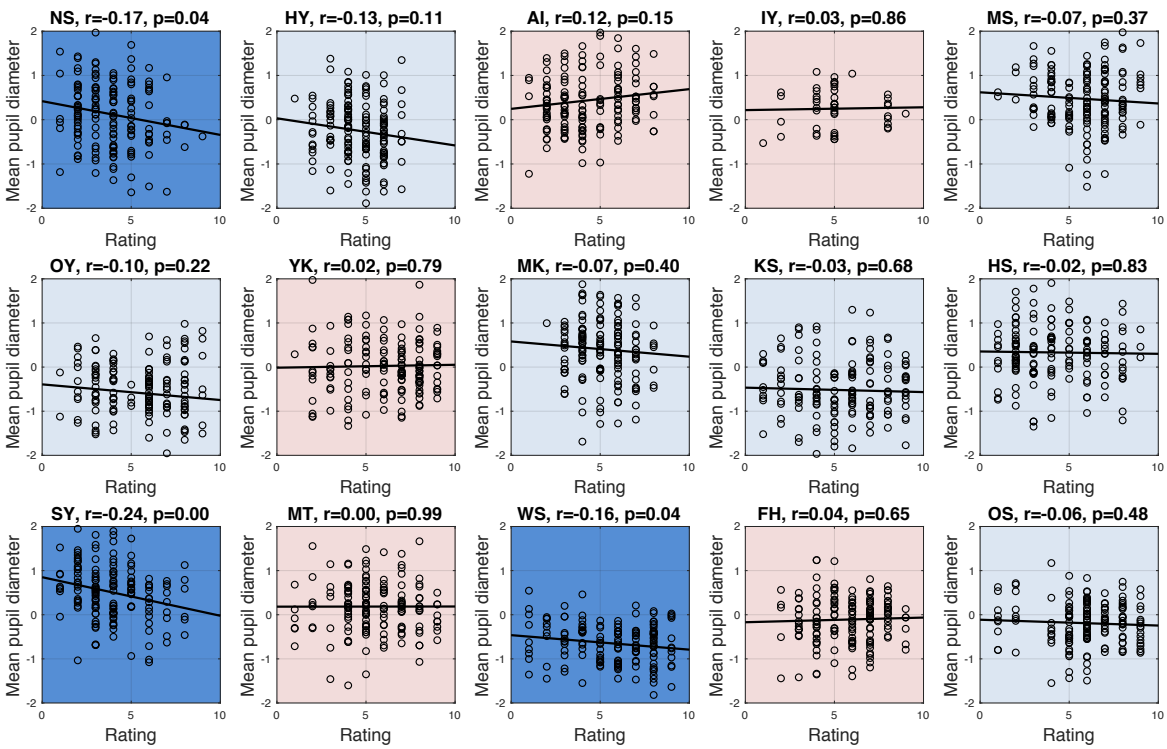
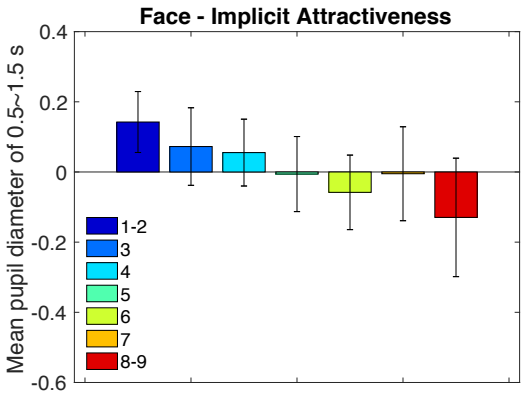


Figure S2C. Supplementary information to Figure 2(C). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the face roundness condition in Experiment 2.

(A) Mean pupil diameter as a function of rating scores. Error bars represent standard errors among participants.



(B) Scatterplots of mean pupil diameter (averaged across the period 0.5~1.5 s after stimulus onset) as a function of implicit attractiveness rating for faces in Experiment 2. Each plot represents each participant. Each dot represents response to each face. Plots in blue represent negative correlation, and the darker blue represents statistical significance of a negative Pearson's correlation. Plots in pink represent positive correlation, but it did not reach statistical significance. For this condition, 13 out of 15 participants showed a tendency of negative correlation between pupil size and implicit attractiveness rating, among which four reached statistical significance individually.

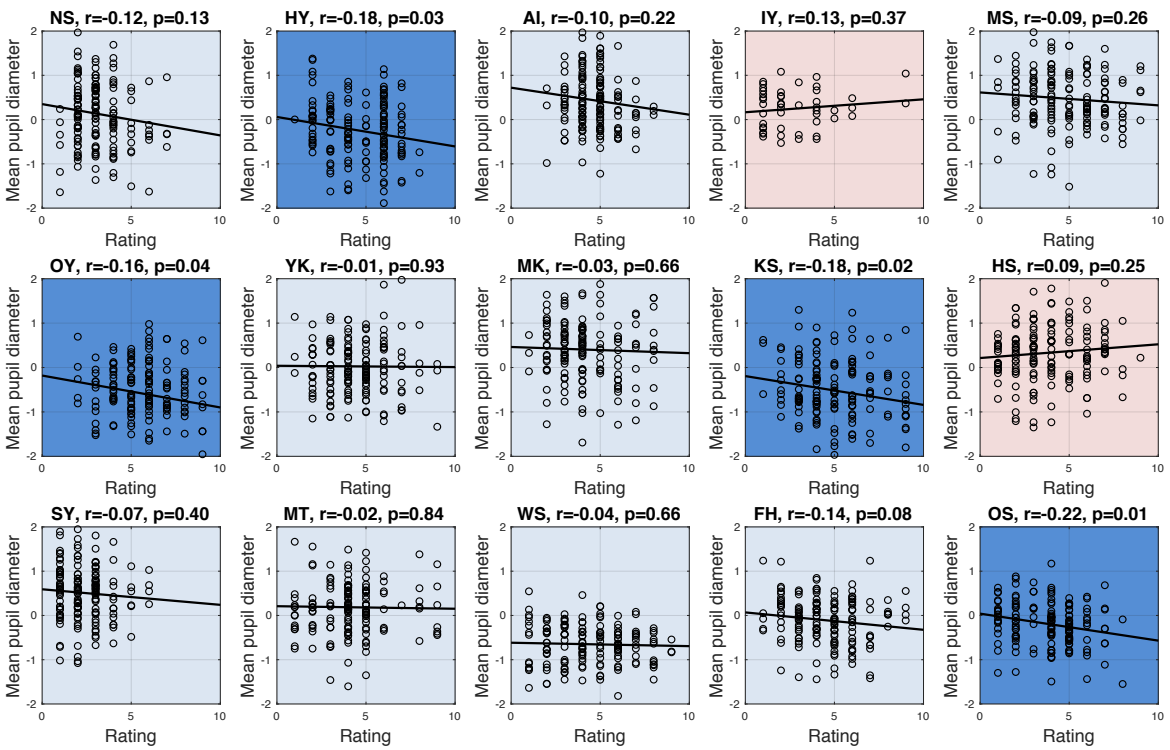


Figure S2D. Supplementary information to Figure 2(D). Plots for (A) ANOVA and linear trend analysis of the mean pupil diameter and (B) correlation analysis for individual participants in the implicit facial attractiveness condition in Experiment 2.

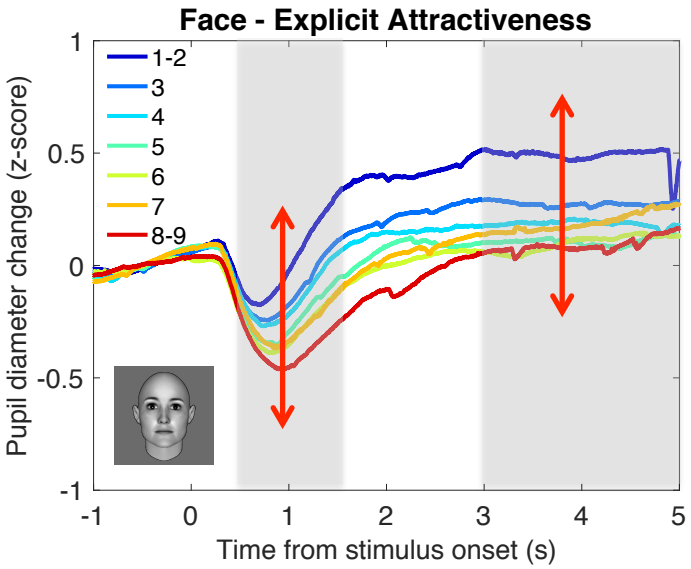


Figure SE1: Pupil response results in Supplementary Experiment 1. Mean pupil diameter as a function of time reference to the target onset during attractiveness judgment for faces. Curves are parameterized with average rating across participants (1 for least attractive and 9 for most attractive). The gray shadow represents the time window for averaging the pupil size to present the amount of pupil response for statistical analysis (see Supplementary Table 1).

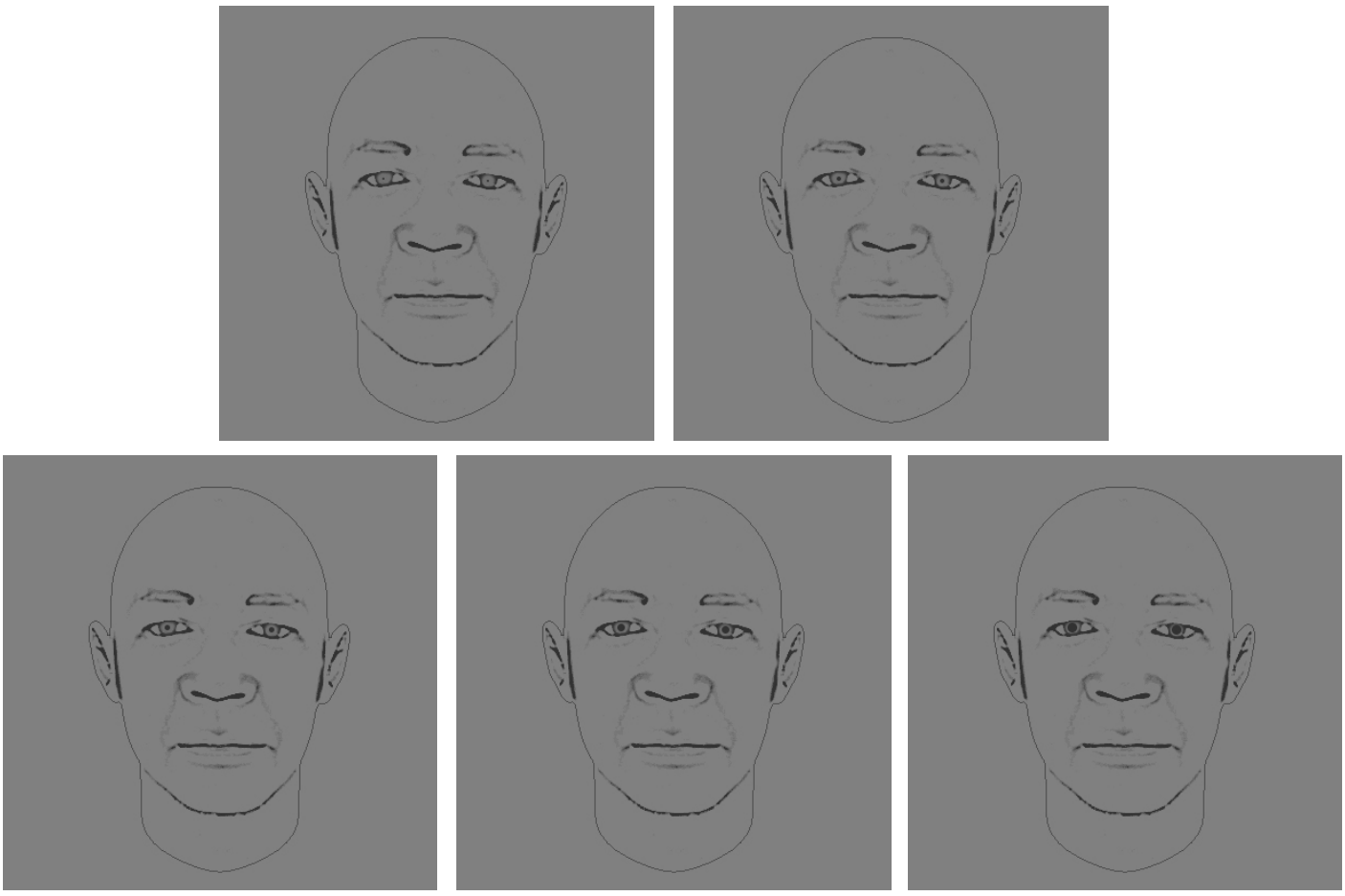


Figure SE2-1: Sample stimuli used in Supplementary Experiment 2

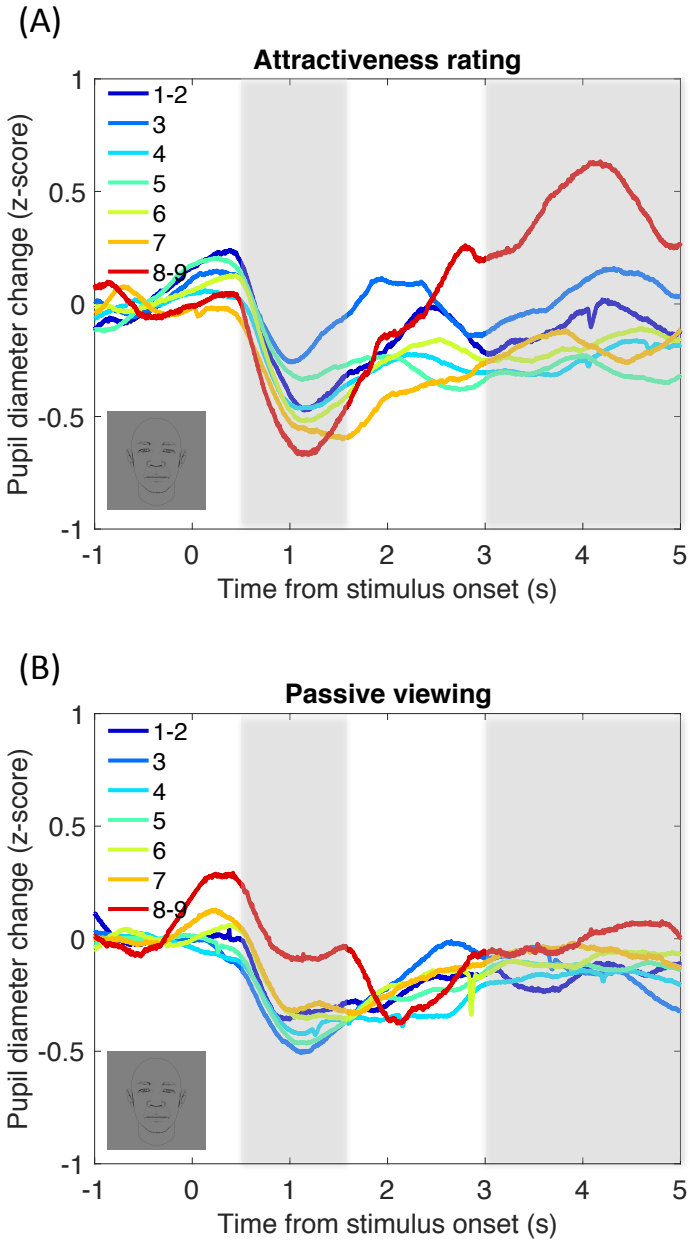


Figure SE2-2: Pupil response results in Supplementary Experiment 2. Mean pupil diameter as a function of time reference to the target onset during (A) attractiveness judgment for faces or (B) passive viewing. Curves are parameterized with average rating across participants (1 for least attractive and 9 for most attractive). The gray shadow represents the time window for averaging the pupil size to present the amount of pupil response for statistical analysis (see Supplementary Table 1).

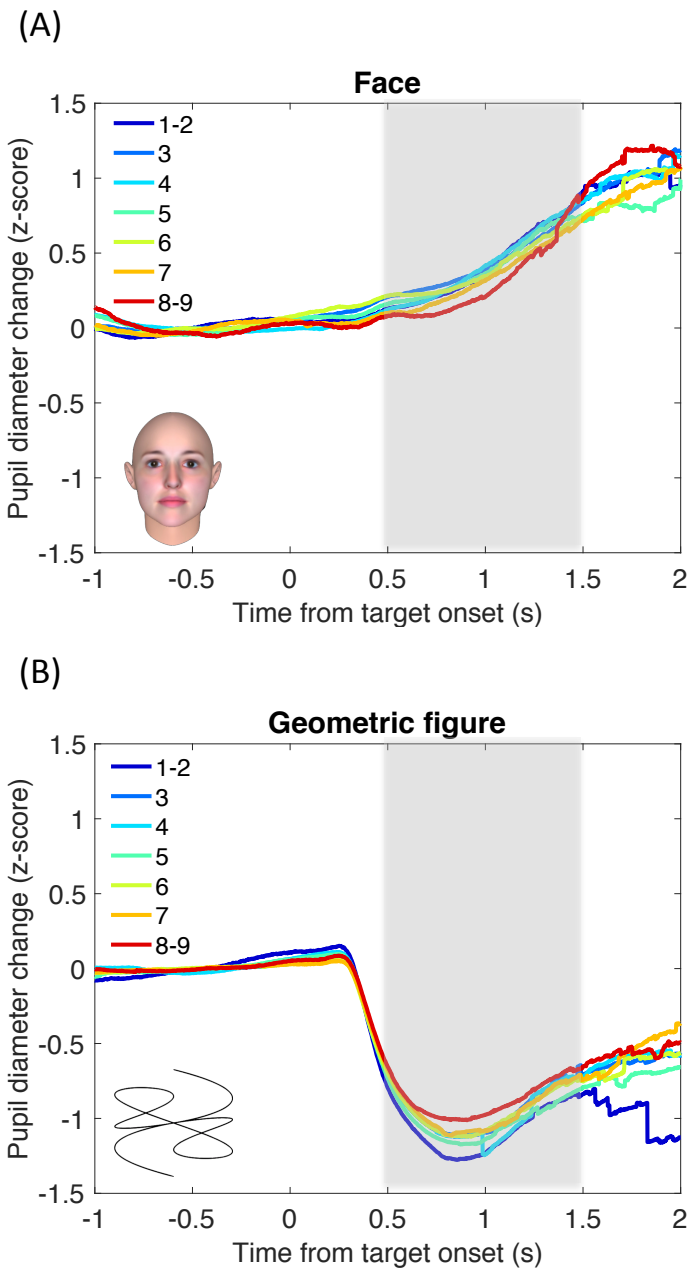


Figure SE3: Pupil response results in Supplementary Experiment 3. Mean pupil diameter as a function of time reference to the target onset during attractiveness judgment for (A) faces or (B) geometric figures. Curves are parameterized with average rating across participants (1 for least attractive and 9 for most attractive). The gray shadow represents the time window for averaging the pupil size to present the amount of pupil response for statistical analysis (see Supplementary Table 1).

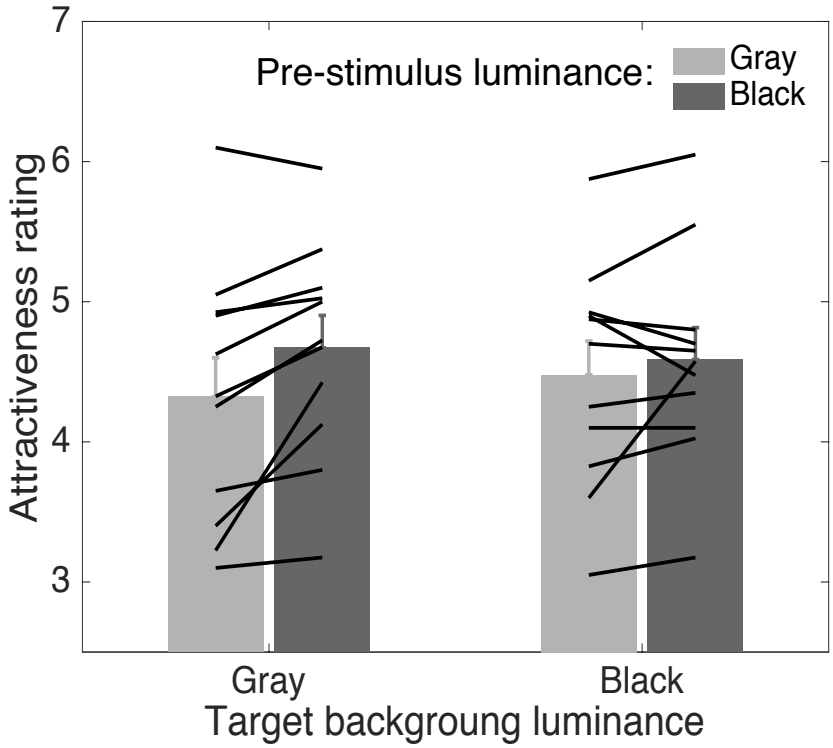


Figure S3. Supplementary information to Figure 3. Mean attractiveness rating as a function of pre-stimulus and target background luminance in Experiment 3. Each black line represents individual participants' mean rating score. Error bars represent standard errors among participants. As shown in the figure, faces were rated more attractive following the gray than black pre-stimulus background, regardless of target background luminance.

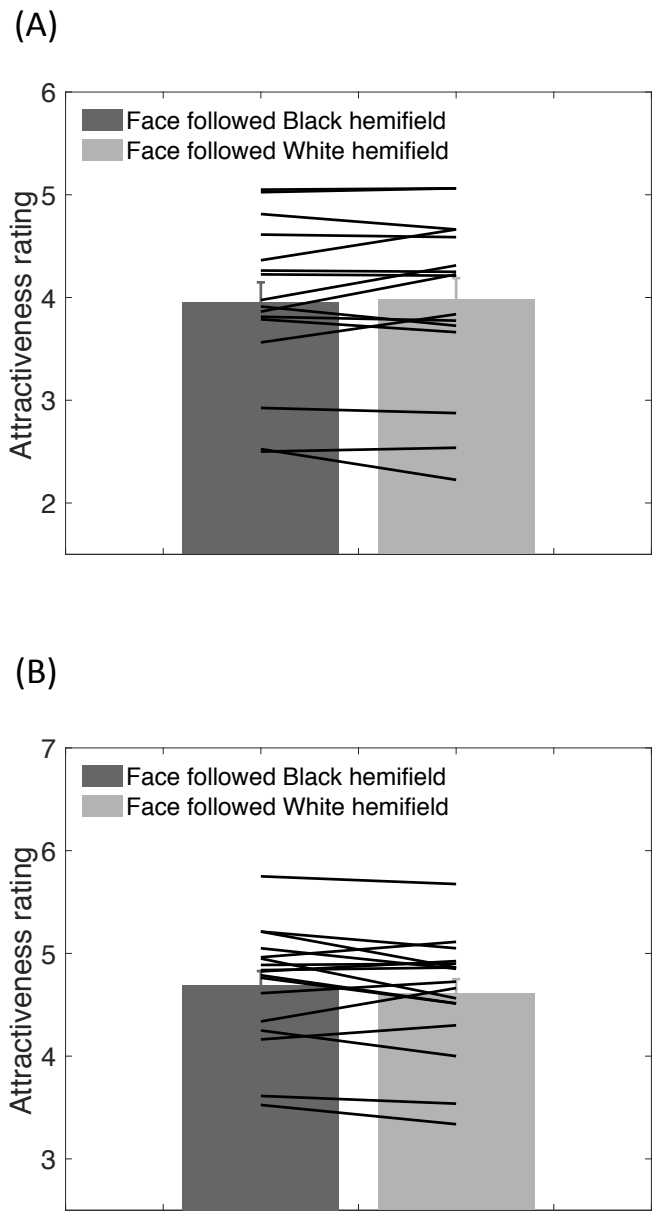


Figure S4-1. Supplementary information to Figure 4. Mean attractiveness rating as a function of pre-stimulus's hemifield's luminance in Experiment 4 when (A) eye movement was allowed or (B) eye movement was not allowed (participants fixated the central fixation cross throughout the trial). Each black line represents individual participants' mean rating score. Error bars represent standard errors among participants.

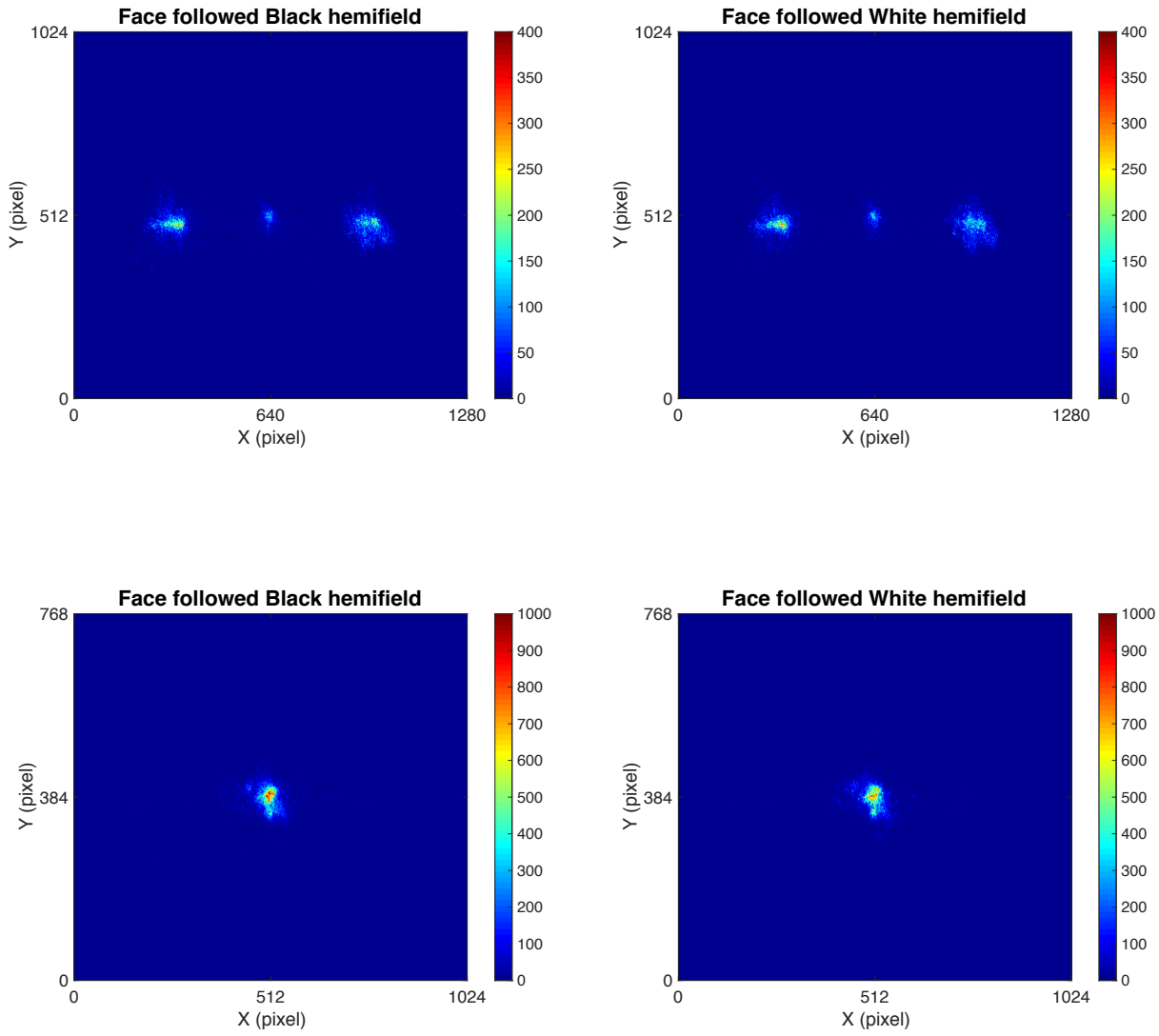


Figure S4-2. Heat map of gaze distribution 0-2 s after the target onset in Experiment 4 when (A) eye movement was allowed or (B) eye movement was not allowed (participants fixated the central fixation cross throughout the trial)